

## Durham Research Online

---

### Deposited in DRO:

11 August 2016

### Version of attached file:

Accepted Version

### Peer-review status of attached file:

Peer-reviewed

### Citation for published item:

Gorard, S. and Makopolou, K. (2012) 'Is mixed methods the natural approach to research?', in Research methods in physical education and youth sport. Abingdon, Oxon: Routledge, pp. 106-119.

### Further information on publisher's website:

<http://www.routledge.com/9780415618854>

### Publisher's copyright statement:

This is an Accepted Manuscript of a book chapter published by Routledge in Research Methods in Physical Education and Youth Sport on 30/11/2011, available online: <http://www.routledge.com/9780415618854>

### Additional information:

---

### Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in DRO
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full DRO policy](#) for further details.

## **CHAPTER 9**

### **Are mixed methods the natural approach to research?**

Stephen Gorard and Kyriaki Makopoulou, University of Birmingham, UK

The authors are pleased to announce the death of mixed-methods research.

## **INTRODUCTION**

In this chapter we argue that mixed methods are so obviously the natural approach to research that we have no real need of the epithet ‘mixed’. If researchers do, or should, naturally use whatever methods they need to answer their research questions, then there is no methods schism. This means that there are no separate elements to be ‘mixed’.

Stephen Gorard has been saying this, or something like it, for a decade or more. Kyriaki Makopoulou wanted to know why. This chapter tries to capture some of the discussion between them in the hope that this will provoke similar discussions for readers.

## **ABOUT US**

Stephen Gorard is an established researcher in education who routinely uses a range of designs including trials, comparative, longitudinal and case study, with a range of techniques for data collection and analysis including historical archives, documents, focus groups, a variety of interviews, participant observation, surveys, modelling, secondary sources and meta-analysis. He uses every one of these with every other, and has done so since completing his PhD in 1996 in response to the variety of projects on which he has worked. He has never believed in the artificial separation of techniques into qualitative and quantitative work, finding that this acts as a barrier to

understanding. His recent and current research includes investigation into the links between family behaviour, poverty and success at school, the socio-economic determinants of participation in science subjects, and the role of enjoyment in learning. He also publishes regularly on the conduct and quality of social science research.

Kyriaki Makopoulou is an emergent researcher in sport pedagogy. Kyriaki joined the School of Education at the University of Birmingham in September 2008, having been a research associate at Loughborough University since 2005. She was trained in the University of Athens between 1996 and 2000, in a faculty with a powerful orientation towards quantitative, experimental research designs. Research methods modules were designed with the intention of training students to develop objective data collection tools and, upon graduation, she was, perhaps naively, convinced that good research derives from a good hypothesis. She discovered qualitative research while studying for a master's degree in the UK, and was fascinated by the power of personal experiences. Drawing upon qualitative methodologies, her recent research focused on the nature of physical education (PE) teachers' engagement in career-long professional learning and its impact upon practice.

## **WHAT IS MIXED-METHODS RESEARCH AND WHEN MIGHT IT BE APPROPRIATE?**

Stephen Gorard poses the question: 'What does it mean to mix methods?' The chapters in this book so far imply that what is being mixed are 'qualitative' and 'quantitative' approaches, although there is of course a range of different kinds of methods that could be mixed (Symonds and Gorard, 2010). Mixing visual and oral datasets is mixing

methods, for example, and it is not clear why everything involving numbers is counted as one approach, and everything else including smells, drawings, acting, music and so on is treated as an alternate monolith called 'qualitative'. Yet even if we use this very simple idea of just two 'Q' categories, it can still be argued that mixing methods is clearly the natural approach in social science research. It is so natural that for me it does not even seem like there is any mixing, because I do not separate the qualitative and quantitative approaches in the first place. In fact, I would argue that this natural synthesis of different kinds of data is what we all do in our everyday lives whenever we are faced with a task with an outcome that is important to us.

Education is an important applied field and the results of research, if taken seriously, can affect the lives of real people and lead to genuine expenditure and opportunity costs. It is instructive to contrast how we, as researchers, sometimes behave when conducting research professionally with the ways we behave when trying to answer important questions in our personal lives. When we make real-life decisions about where to live, where to work, the care and safety of our children and so on, most of us behave very differently from the way we do as 'researchers'. If, for example, we were intending to purchase a house by paying most of our savings and taking out a mortgage for 25 years that is equal in size to many times our salary, then we would rightly be cautious. We would have many crucial questions to answer from the beginning, and would only go ahead with the transaction once assured that we had sufficiently good answers from what is, in effect, a serious piece of research. It is worth considering this example in some detail because it illustrates fundamental issues in a very accessible way.

When purchasing a house, we will believe that the house is real even though external to us, and that it remains the same even when we approach it from different ends of the street (else why would we buy it?). Thus, we would not start with 'isms' or paradigms. We would not refuse to visit the house, or talk to the neighbours about it, because we were 'quantitative' researchers and did not believe that observation or narratives were valid or reliable enough for our purposes. We would not refuse to consider the interest rate for the loan, or the size of the monthly repayments, because we were 'qualitative' researchers and did not believe that numbers could do justice to the social world. In other words, in matters that are important to us personally, there is a tendency to behave logically, eclectically, critically, and sceptically. We collect all and any evidence available to us as time and resources allow, and then synthesize it quite naturally and without considering mixing methods as such. I have long argued that academic research should be considered in the same way. For me, this means no Q words, no paradigms and no isms.

I do not believe that types of data or methods of data collection and analysis have paradigmatic characteristics, and so for me there is no problem in using numbers, text, visual and sensory data synthetically (Gorard, 2010a). The methods of analysis for text, numbers and sensory data are largely the same, consisting of searching for patterns and differences, establishing their superficial validity and then trying to explain them. Other commentators and methods resources may try to claim that there is a fundamental difference between looking for a pattern or difference in some measurements and in some text or observations. This unnecessarily complex view is based on a number of widely held logical fallacies that get passed on to new researchers under the guise of

research methods training (Gorard, 2010b). There are no 'qualitative' and 'quantitative' paradigms; working with numbers does not mean holding a view of human nature and knowledge that is different from when you work with text. As noted earlier, my position is that the whole schism is nonsense, and so mixing is not needed since there is nothing separate to mix.

In the sociology of science, the notion of a 'paradigm' is a description of the sets of socially accepted assumptions that tend to appear in 'normal science' (Kuhn, 1970). A paradigm is a set of accepted rules within any field for solving one or more puzzles – where a puzzle is defined as a scientific question to which it is possible to find a solution in the near future. An example would be Newton setting out to explain Kepler's discoveries about the motions of the planets. Newton knew the parameters of the puzzle and so was working within a paradigm. A more recent example might be the Human Genome Project, solving a closely defined problem with a widely accepted set of pre-existing techniques. Such puzzles can be distinguished from the many important and interesting questions that do not have an answer at a particular stage of progress (Davis, 1994). The 'normal science' of puzzles in Kuhnian terms is held together, rightly or wrongly, by the norms of reviewing and acceptance that work within that taken-for-granted theoretical framework. A paradigm shift occurs when that framework changes, perhaps through the accumulation of evidence, perhaps due to a genuinely new idea, but partly through a change in general acceptance. Often a new paradigm emerges because a procedure or set of rules has been created for converting another more general query *into* a puzzle.

Yet, instead of using 'paradigm' to refer to a topic or field of research (such as traditional physics) which might undergo a radical shift on the basis of evidence (to quantum physics, for example), some commentators now use it to refer to a whole approach to research including philosophy, values and method (Perlesz and Lindsay, 2003). The most common of these approaches are qualitative and quantitative, even though the Q words only make sense, if they make sense at all, as descriptions of data. It could be argued that these commentators use the term 'paradigm' to defend themselves against the need to change, or against contradictory evidence of a different nature to their own. In such cases, the idea of paradigm functions to defend them because they (pointlessly) parcel up unrelated ideas in methodology. Thus, the idea of normal science as a collection of individuals all working towards the solution of a closely defined problem has all but disappeared in the social sciences. Instead, we have paradigm as a symptom of scientific immaturity.

The result of a defensive use of the term 'paradigm' is that the concept has become a cultural cliché with so many meanings it is now almost meaningless. Many of the terms associated with paradigms – i.e. the isms such as positivism – are used almost entirely to refer to others, having become intellectually acceptable terms of abuse and ridicule (see also Hammersley, 2005). Yet, surely most of us could agree that 'Research should be judged by the quality and soundness of its conception, implementation and description, not by the genre within which it is conducted' (Paul and Marfo, 2001: 543–5). The paradigm argument that reinforces the differences between the Q word approaches is a red herring and, I would argue, unnecessarily complex.

One common argument for difference between the Q word approaches is their scale (e.g. Creswell and Plano Clark, 2007). It is argued that qualitative data collection necessarily involves small numbers of cases, whereas quantitative relies on very large samples in order to increase power and reduce the standard error. Yet even this is misleading for two main reasons. First, it is not an accurate description of what happens in practice. The accounts of hundreds of interviewees can be properly analysed as text, and the account of one case study can properly involve numbers. The supposed link between scale and paradigm is just an illusion. Second, issues such as sampling error and power relate to only a tiny minority of quantitative studies where a true and complete random sample is used or where a population is randomly allocated to treatment groups. In the much more common situations of working with incomplete samples with measurement error or dropout, convenience, snowball and other non-random samples and the increasing amount of population data available to us, the constraints of sampling theory are simply not relevant. It is also the case, as I have argued elsewhere, that the standard error/power theory of analysis is fatally flawed in its own terms, even when used as intended (Gorard, 2010b). It is based on the logical error of mistaking the probability of encountering the data observed given a true hypothesis (for example, what significance tests calculate) for the probability of the hypothesis being true given the data observed (what significance testers actually want, and usually pretend that they have calculated).

Qualitative research, so its proponents argue, is supposed to be subjective and thus closer to a social world (Gergen and Gergen, 2000). Quantitative research, on the other hand, is supposed to help us become objective (Bradley and Schaefer, 1998). This



distinction between quantitative and qualitative analysis is exaggerated, largely because of widespread error by those who do handle numbers (Gorard, 2010b) and ignorance of the subjective and interpretivist nature of numeric analysis by those who do not (Gorard, 2006). What few seem to recognize is that the similarities in the underlying procedures used are remarkable (Onwuegbuzie and Leech, 2005). Few analytical techniques are restricted by data gathering methods, input data, or by sample size. Most methods of analysis use some form of number, such as ‘tend’, ‘most’, ‘some’, ‘all’, ‘none’, ‘few’ and so on (Gorard, 1997). Whenever one talks of things being ‘rare’, ‘typical’, ‘great’ or ‘related’, this is a numeric claim, and can only be so substantiated, whether expressed verbally or in figures (Meehl, 1998). Similarly, quantification does not consist of simply assigning numbers to things, but of linking empirical relations to numeric relations (Nash, 2002). Personal judgements lie at the heart of all research – in our choice of research questions, samples, questions to participants and methods of analysis – regardless of the kinds of data to be collected. The idea that quantitative work is objective and qualitative is subjective is based on a misunderstanding of how research is actually conducted.

## **RESPONSE FROM AN EMERGENT RESEARCHER**

Kyriaki Makopoulou writes: I would like to probe the ‘objective’ versus ‘subjective’ issue further as there is a lot of confusion surrounding it. Giacobbi *et al.* (2005: 23) write:

While mixing methods from different paradigms is possible, desirable, and often productive, the underlying assumptions of various paradigms (i.e. constructivists

versus positivism) *may contradict one another* (Krane and Baird, in press; Lincoln and Guba, 2000). In other words, a constructivist may use quantitative data but will adopt a subjective epistemology, while a positivist who uses a post-experiment interview will do so under an objective epistemology.

Debates on objective versus subjective research are also implicated in discussions about research quality. Quantitative researchers consider concepts like validity, reliability, generalizability and objectivity to be essential criteria in assessing research quality (Healy and Perry, 2000). From a qualitative perspective, these terms are not always adequate to encapsulate the range of issues that affect quality in qualitative research (Seale, 1999) and new concepts and criteria have been developed. For example, it is claimed that qualitative research should be ‘trustworthy’, replacing the notion of research validity. When the definitions of these apparently different terms are unpacked, it is clear that they share (more or less) the same meaning. From the perspective of those who champion different paradigms they do differ, however, in the degree of claimed neutrality.

Researchers adopting a positivistic approach aim to consciously avoid personal involvement that might bias a study. This means that researchers try to detach themselves from the ‘object’ of inquiry – in essence, depersonalize their research – with the aim of capturing and communicating the ‘truth’ in an ‘objective’ manner. Schwandt (2000) defined ‘bias’ as the tendency for researchers to impose a priori theoretical frameworks or interpretations on the data, marginalize or exclude opposite or alternative perspectives, and draw unjustifiable inferences or generalizations. Many qualitative

researchers, although acknowledging that any factors influencing findings in an unethical way should be eliminated, accept a degree of involvement as an inevitable part of the inquiry and treat claimed neutrality as ‘deluding’ or ‘misleading’. Greenbank (2003) captured this as the debate between those advocating a value-neutral and those arguing for a value-laden approach to (educational) research. Seale (1999: 3) talked about a clash of two moments.

This debate is prominent and persisting, although I also support this chapter’s position that researchers should not ‘tie’ themselves to a particular paradigm. The key argument is that dualistic thinking (either–or) is problematic in an ever-evolving world that necessitates multidimensional insights and solutions to complex problems. For decades, it has been acknowledged that a rigid, unreflective adherence to a research paradigm might encourage researchers to take fundamental assumptions for granted (Patton, 1978) thereby preventing them from learning from other researchers working within apparently contrasting research perspectives (Bailey, 2007). This chapter’s position is that engaging in such a discussion in the first place is restrictive. Should researchers in their search for ‘warranted’ evidence (Gorard, 2002) therefore strive for neutrality in the process or should they acknowledge a degree of involvement? Or should debates about the degree of neutrality be abandoned altogether and replaced by discussions on research designs that are robust, rigorous and transparent? If so, what would such a research design look like?

There is another apparent constraint in endeavours to ‘mix’ methods. In most cases, quantitative studies (especially experimental designs) begin with a *research hypothesis*,

which is the expected result (Thomas *et al.*, 2010), and which is tested to be supported or refuted. Qualitative studies, on the other hand, set out to explore answers to *research questions*. In most forms of research, both hypothesis and questions should be developed following a thorough, in-depth review of the available literature. The logical question that follows from this, however, is: What is the starting point of a study that is not framed by the ‘paradigmatic war’? Should researchers pose questions and/or hypotheses?

### **SOME IDEAS ON HOW TO CONDUCT MIXED-METHODS STUDIES**

Stephen Gorard responds to the questions posed by Kyriaki Makopoulou: If we consider just some of the ways in which methods can be mixed within one study or programme it becomes obvious, to me, that these questions should be reconsidered. A programme of research conducted by one team, or a field of research conducted by otherwise separate teams, will incorporate most methods of data collection and analysis. Figure 9.1 is a simplified description of a full cycle for a research programme. The cycle is more properly a spiral which has no clear beginning or end and in which activities (phases) overlap, can take place simultaneously, and iterate. Starting with draft research questions, the research cycle might continue with a synthesis of existing evidence (phase 1).

Figure 9.1 near here

Figure 9.1: An outline of the full cycle of social science research and development

Ideally this synthesis would be an inclusive review of the literature both published and unpublished, coupled with a re-analysis of relevant existing datasets of all kinds (including data archives and administrative datasets), and related policy/practice documents. It is impossible to conduct a fair appraisal of the existing evidence on almost any topic in applied social science without drawing upon evidence involving text, numbers, pictures and a variety of other data forms. Anyone who claims to be conducting even the most basic literature review without combining numeric and textual data is surely misguided. One way (of many) to combine disparate datasets is via Bayesian synthesis, where the in-depth evidence and perhaps professional judgements provide a subjective a priori probability for a research claim, with numeric data used to adjust this base a posteriori (Gorard *et al.*, 2004). It can be argued that this relatively straightforward method of generating a result is fair and inclusive, and requires little technical expertise. It also illustrates that, by definition, the outcomes will be subjective. Indeed, no matter how transparent and technically sound a process of analysis might be, the result will always be subjective for the reasons Kyriaki Makopoulou suggests above. We might argue, therefore, that conducting this kind of synthesis is one way to persuade those who purport to be ‘quantitative’ researchers that their work is not quite as objective as they claim, even when only a single method is used.

Where a project or programme continues past phase 1 in Figure 9.1, every further phase in the cycle tends to require a mixture of methods. The overall research programme might be envisaged as tending towards an artefact or ‘product’ of some kind. This product might be a theory (if the desired outcome is simply knowledge), a proposed improvement for public policy, or a tool/resource for a practitioner. In order for any of

these outcomes to be promoted and disseminated in an ethical manner, however, they must have been *tested*. If there is no testing, the dissemination can only state (ethically) that the product or output seems to be a good idea but that we have no real idea of its value. A theory, by definition, will generate testable propositions. A proposed public policy intervention can be tested realistically and then monitored in situ for the predicted benefits, and for any unwanted and undesirable side effects. Therefore, for the minority of programmes that continue as far as phase 6 in Figure 9.1, rigorous testing would usually involve a mixture of methods and types of evidence in just the same way as phase 1.

Even where a purely numeric outcome is envisaged as the benefit of the research programme (such as a more effective or cost-efficient service) it is pointless knowing that the intervention works if we are not also aware that, for example, it is unpopular and therefore likely to be ignored or subverted in practice. Similarly, it would be a waste of resource, and therefore unethical, simply to discover that an intervention did not work in phase 6 and so return to a new programme of study in phase 1. We would want to know *why* it did not work, or perhaps how to improve it, and whether it was effective for some regular pattern of cases but not for others. So in phase 6, as in phase 1, the researcher or team who genuinely wants to find something out will, quite naturally, use a range of methods and approaches including measurement, narrative and observation.

The key point to be made here is that a research design, such as case study, longitudinal research or randomized controlled trial is completely independent of individual methods

of data collection and analysis. Research design, including the pre-specification of success and failure criteria, is one important defence against the inevitable researcher bias in any study. Replication, syntheses and critical judgements are others. The best examples of each design are likely to use all and any available data collection methods irrespective of type and without consideration of invented epistemological schisms.

A range of simple ways in which data can be mixed are given in Gorard with Taylor (2004). One of these suggested ways is termed ‘new political arithmetic’ but is really just a sequence of related questions answered by different kinds of data. Typically, large-scale data (perhaps already existing from official sources) is used to define a problem, pattern, trend or difference. It is also used to select a representative subset of cases for in-depth research to investigate the reasons for the problem, pattern, trend or difference. The in-depth work is, therefore, generalizable in the sense that this term is traditionally used, and different datasets are used to define the pattern and its determinants. Again what an example such as this makes clear is that all research answers a question, so the notion of a formal hypothesis is not necessary to describe the process. It follows, therefore, that the concern raised earlier by Kyriaki Makopoulou – hypothesis or research question – can be addressed quite easily. A question can be converted to a hypothesis simply by assuming one answer to the question, and vice versa, so that a hypothesis immediately generates the question ‘Is it true?’ For example, the question ‘Are men and women different in this regard?’ is the same as the hypothesis ‘Men and women are different in this regard,’ which can be true or false. Any claim about a fundamental difference between the two, as made by Thomas *et al.* (2010) for example and cited earlier, is just plain wrong.

## **SOME EXAMPLES OF REAL-LIFE STUDIES**

What happens currently in practice in sport pedagogy research? As part of the preparation for this chapter, Kyriaki Makopoulou hand-searched the articles in two journals – *Sport, Education and Society* and *Physical Education and Sport Pedagogy* – over three years (2008 to 2010). Both are based in the UK but international in remit and content, and we use these to give us an idea of the range of methods and types of work being conducted in the field. The journals contained 135 peer-reviewed articles, of which 28 (21 per cent) had no empirical evidence. This is rather less than the proportion of ‘thought pieces’ we have found in more generic reviews (e.g. Gorard *et al.*, 2007). Of the remainder, 72 (67 per cent) were described as ‘qualitative’ and made no use of numeric evidence at all. Therefore in this admittedly limited sample, over two-thirds of research involves no numbers. (It should also be acknowledged that some other journals in the field have a bias towards publishing quantitative research.)

The preponderance of qualitative research in the two journals selected, both of which have a social science base, is in line with more generic reviews both in education and most social sciences more widely. It is astonishing but apparently true. Around 21 per cent (23 papers) used numeric evidence only and these were largely based on sampling theory derivatives (often incorrectly). This leaves only 12 papers (11 per cent) that used both numeric and some other form of evidence, usually interviews. What this suggests is that ‘mixed-methods’ papers, if that is what these were, are a minority at least for the authors, editors and readers of these two journals. Similar studies have found even smaller proportions. Examining the methodologies employed in published dissertations



in the United States, Silverman and Manson (2003: 291) found that a combination of both qualitative and quantitative methods was used in 8 per cent of dissertations. Similarly, Ward and Ko (2006) investigated articles published in the *Journal of Teaching in Physical Education* by methodology from 1981 to 2005 and reported that of the 68 per cent of articles that were research studies, only 6 per cent were mixed methods. Likewise, a review conducted in 2003 found that out of 485 published papers in three leading sport psychology journals, only 5 per cent adopted a mixed-methods approach (Giacobbi *et al.*, 2005).

Two decades ago, Schempp (1989) identified a strong research tradition influenced by the natural and biological sciences in the PE and sport pedagogy research field. If it were true then, it is clearly not so now. For some time, the number of studies drawing upon so-called 'qualitative' methodologies has dominated output. Meanwhile, qualitative and quantitative methodologies are still viewed as two fundamentally distinct approaches, perhaps appropriate for answering different kinds of research questions but underpinned by diverse philosophical and epistemological standpoints. This means that in the few examples of apparent mixed-methods work found in our search of the journals, qualitative and quantitative methodologies are not really being 'mixed' in the field of sport and sport pedagogy.

A clear example is the study by Lee *et al.* (2007) from the sport psychology field. This drew upon five other studies to develop and validate a questionnaire, but the paper is a report of a survey study with a small focus group thrown in, almost as an addition. It displays many of the common weaknesses of standard quantitative research. The data

they use for the factor analysis are not interval in nature. The factor analytic model explains very little of the variance in the dataset, and the majority of variance is just ignored in the analysis, with no reason given. The authors just omit uncorrelated items (but see Gorard, 2010c). The use of significance tests is completely unjustified as there is no randomization involved, and so the probabilities they quote are meaningless but also misleading. A further example, Morgan and Hansen (2008), is very similar. The different kinds of data are dealt with separately rather than synthesized, and numeric analysis predominates but is of the usual poor quality. For example, their model only explains 32 per cent of the variance, they imply an unjustified causal model from mere association, and they conduct their significance tests at the level of individuals when the only randomization is of schools. If these examples are really mixed methods rather than just two isolated streams of data dealt with in the same report, then they do not appear very natural (or very good). There is a way to go yet, before evidence outweighs paradigmatic clutter.

Kyriaki Makopoulou, meantime, is developing a new research proposal on continuing professional development (CPD) for PE teachers (PE-CPD). Her intention is that with guidance from colleagues, including Stephen Gorard, it *will* be a mixed-methods study in the full sense of the term. The research cycle began from phase 1, by drafting a set of research questions to guide an ‘inclusive’ synthesis of available evidence:

- What is currently known about CPD practices that have positive impact on teachers and pupils? What needs to be known to inform CPD policy and practice?

- What kind of evidence can provide a robust basis for CPD policy and practice?

The synthesis suggested that a significant shift in CPD research is required: towards experimental or quasi-experimental research designs that measure impact on pupils as well as teachers. However, it was also noted that recent efforts to test the relative effectiveness of different CPD types yielded disappointing results (Garet *et al.*, 2008). Even more problematically, collecting numerical data alone in the form of pre- and post-test measurements provided disappointing findings as researchers were unable to offer insights into how and why the interventions failed to work. In other words, as Stephen Gorard pointed out earlier, these studies lack explanatory power for practice.

The analysis undertaken in preparation for the proposed PE-CPD study (including the review of the relevant literature and policy documents) showed that in order to advance scientific knowledge and to impact CPD policy and practice, a mixed-methods approach to CPD research – one that is advocated as ‘natural’ in this chapter – is required. To address this gap (and moving into subsequent phases of Stephen Gorard’s cycle suggested above), the proposed study will employ an experimental design and use mixed methods to investigate the relative effectiveness of two different forms of CPD for both teacher and pupil learning. In particular, a CPD programme will be designed, delivered and evaluated rigorously to assess impact on teachers and pupils. The proposed study will employ a three-arm experimental design. One treatment will be the usual three-day block workshops. A second treatment will involve the same amount of involvement in CPD but with shorter meetings spread across a longer period of time. The third group will be a control with no treatment.

The next phase in the research cycle is to conduct a feasibility study, involving a mixture of methods. This is a vital and ethical practice to ensure, before implementing the CPD intervention with the wider PE population, that we can be sure how, why, when, for whom and under what circumstances it works. As Stephen Gorard stressed above, it would be ‘unethical ... to discover that an intervention did not work in phase 6 and so return to a new programme of study in phase 1’. In the CPD literature, feasibility studies are defined as studies that measure the impact of a CPD intervention delivered by the same provider in one area. The proposed CPD intervention will be implemented in one local authority in England. A thorough investigation of the PE teachers is required to ensure that the PE-CPD intervention has meaning and relevance to the participant teachers, and can be tailored (personalized) to their needs and expectations.

In all subsequent phases of the project, a mixture of methods will be employed to ensure that the research team has a holistic picture of the nature and impact of the CPD intervention. For example, to measure impact on pupil learning, a traditional achievement test will be developed to measure pupil learning in terms of health-related components and aligned with the expectations outlined in the attainment section of the national curriculum for physical education. However, it is also important to explore pupils’ experiences of health-related exercise in PE and explore the ways pupils construct meanings and how they believe that their experiences will influence their participation in physical activity.

## **CONCLUSIONS**

An analysis of papers published in two social science-based journals in the field of physical education and sport pedagogy suggests that researchers do not routinely adopt mixed-methods approaches nor synthesize data of different types. The tendency to publish quantitative and qualitative research undertaken in this field in different journals must surely reinforce the notion of two researcher ‘camps’, each intrinsically hostile to the approach of the other but often without the skills to be appropriately and helpfully critical for each other. This might be one explanation for the finding that much so-called qualitative research tends to lack the rigour found in the best examples, and too much so-called quantitative research is needlessly complex and prone to serious error. Combining approaches, therefore, as long as we combine the best of each approach rather than the worst, holds the promise of providing better answers to important research questions. Recognition of the danger of obscuring research design with polarized paradigmatic clutter could also lead to substantive improvements in single-methods studies.

So what does all this mean in practice for new researchers? We realize that asking new researchers to reject traditional methods teaching and resources is challenging. Perhaps one way to think about this is to ensure that you ask questions about the track record of any person giving methodological advice. You may find that those who argue most strongly for the importance of the Q word paradigms also conduct the weakest research, using an impoverished set of designs and techniques repeatedly. This should, at the very least, raise some questions about their critical understanding of the wider research landscape and the potential for ‘mixing’ methods. So, we argue that if a researcher really cares about finding something out that is as robust as possible, they should

consider ignoring the traditional two-camp research methods resources and behave in research as they would in real life. In real life, the use of mixed methods is natural – so natural, in fact, that we do not generally divide data in the first place. The question to be asked, therefore, is why research should be any different.

## KEY TERMS

Mixed methods, combining methods, research paradigms

## REFERENCES

- Bailey, R. (2007) 'The problem with paradigms', invited paper presented at the 12th Annual Congress of the ECSS, 11–14 July, Juvaskyla, Finland.
- Bradley, W. and Shaefer, K. (1998) *Limitations of Measurement in the Social Sciences*, Thousand Oaks, CA: Sage.
- Creswell, J. and Plano Clark, V. (2007) *Designing and conducting mixed methods research*, London: Sage.
- Davis, J. (1994) 'What's wrong with sociology?', *Sociological Forum*, 9, 2: 179–97.
- Gergen, M. and Gergen, K. (2000) 'Qualitative inquiry, tensions and transformations', in N. Denzin and Y. Lincoln (eds) *The Landscape of Qualitative Research: Theories and Issues*, Thousand Oaks, CA: Sage.
- Giacobbi, P. R., Poczwadowski, A. and Hager, P. (2005) 'A pragmatic research philosophy for sport and exercise psychology', *The Sport Psychologist*, 19: 18–31.
- Gorard, S. (1997) 'A choice of methods: the methodology of choice', *Research in Education*, 57: 45–56.

- Gorard, S. (2002) 'Fostering scepticism: the importance of warranting claims', *Evaluation and Research in Education*, 16, 3: 136–49.
- Gorard, S. (2006) 'Towards a judgement-based statistical analysis', *British Journal of Sociology of Education*, 27, 1: 67–80.
- Gorard, S. (2010a) 'Research design, as independent of methods', in C. Teddlie and A. Tashakkori (eds) *Handbook of Mixed Methods*, Los Angeles: Sage.
- Gorard, S. (2010b) 'All evidence is equal: the flaw in statistical reasoning', *Oxford Review of Education*, 36, 1: 63–77.
- Gorard, S. (2010c) 'Measuring is more than assigning numbers', in G. Walford, E. Tucker and M. Viswanathan (eds) (389–408) *Sage Handbook of Measurement*, Los Angeles: Sage.
- Gorard, S., Roberts, K. and Taylor, C. (2004) 'What kind of creature is a design experiment?', *British Educational Research Journal*, 30, 4: 575–90.
- Gorard, S., with Adnett, N., May, H., Slack, K., Smith, E. and Thomas, L. (2007) *Overcoming barriers to HE*, Stoke-on-Trent: Trentham Books
- Gorard, S. with Taylor, C. (2004) *Combining Methods in Educational and Social Research*, London: Open University Press.
- Greenbank, P. (2003) 'The role of values in educational research: the case for reflexivity', *British Educational Research Journal*, 29(6): 791–802.
- Hammersley, M. (2005) 'Countering the "new orthodoxy" in educational research: a response to Phil Hodgkinson', *British Educational Research Journal*, 31, 2: 139–55.

- Healy, M. and Perry, C. (2000) 'Comprehensive criteria to judge validity and reliability of qualitative research within the realism paradigm', *Qualitative Market Research*, 3(3): 118–26.
- Kuhn, T. (1970) *The Structure of Scientific Revolutions*, Chicago: University of Chicago Press.
- Lee, M. J., Whitehead, J. and Ntoumanis, N. (2007) 'Development of the attitudes to moral decision-making in youth sport questionnaire', *Psychology of Sport and Exercise*, 8: 369–92.
- Meehl, P. (1998) 'The power of quantitative thinking', speech delivered upon receipt of the James McKeen Cattell Fellow award at American Psychological Society, Washington, DC, 23 May.
- Morgan, P. J. and Hansen, V. (2008) 'The relationship between PE biographies and PE teaching practices of classroom teachers', *Sport, Education and Society*, 13(4): 373–91.
- Nash, R. (2002) 'Numbers and narratives: further reflections in the sociology of education', *British Journal of Sociology of Education*, 23, 3: 397–412.
- Onwuegbuzie, A. and Leech, N. (2005) 'Taking the "Q" out of research: teaching research methodology courses without the divide between quantitative and qualitative paradigms', *Quality and Quantity*, 38: 267–96.
- Patton, M. (1978) *Utilization-Focused Evaluations*, London: Sage Publications.
- Paul, J. and Marfo, K. (2001) 'Preparation of educational researchers in philosophical foundations of inquiry', *Review of Educational Research*, 71, 4: 525–47.



- Perlesz, A. and Lindsay, J. (2003) 'Methodological triangulation in researching families: making sense of dissonant data', *International Journal of Social Research Methodology*, 6, 1: 25–40.
- Schempp, P. G. (1989) 'Building the science of teaching in physical education: A crack in the foundation', *Research Quarterly for Exercise and Sport*, 59: 197–9.
- Schwandt, T. D. (2000) 'Three epistemological stances for qualitative inquiry: Interpretivism, hermeneutics, and social constructionism', in N. K. Denzin and Y. S. Lincoln (eds) *The handbook of qualitative research*, 2nd edn, (189–213), Thousand Oaks, CA: Sage.
- Seale, C. (1999) 'Quality in qualitative research', *Qualitative Inquiry*, vol. 5: 465–78.
- Silverman, S. and Manson, M. (2003) 'Research on teaching in physical education doctoral dissertations: a detailed investigation of focus, method and analysis', *Journal of Teaching in Physical Education*, 22: 280–97.
- Symonds, J. and Gorard, S. (2010) 'The death of mixed methods?: or the rebirth of research as craft', *Evaluation and Research in Education*, 23, 2: 121–36.
- Thomas, J. R., Nelson, J. K. and Silverman, S. J. (2010) *Research Methods in Physical Activity*, 5th edn, Champaign, IL: Human Kinetics.
- Ward, P. and Ko, B. (2006) Publication trends in the "Journal of Teaching in Physical Education" from 1981 to 2005, *Journal of Teaching in Physical Education*, 25, 3: 266-280